DOCUMENT RESUME

ED 167 571 **TM 007 681**

AUTHOR	Moore, R. P.; Shah, B. V.
TITLE	Efficiency Study of the NLS Sample Design. 220-884-3.
INSTITUTION	Research Triangle Inst., Durham, N.C. Statistics Research Div.
PUB DATE	Oct 75
NOTE	39p.; Some pages marginally legible due to print quality
EDRS PRICE	MF-$0.83 HC-$2.06 Plus Postage.
DESCRIPTORS	Changing Attitudes; Cost Effectiveness; *Data Analysis; *Efficiency; Followup Studies; Graduate Surveys; High School Graduates; *National Surveys; Participant Characteristics; Predictor Variables; *Research Design; *Sampling; Senior High Schools
IDENTIFIERS	*National Longitudinal Study High School Class 1972; *Sample Size

ABSTRACT

An efficiency study was conducted of the sampling design used for the National Longitudinal Study of the High School Class of 1972 (NLS), and the change variables used in the first followup. Nine alternative designs were compared, using a cost model based upon the number of schools and students sampled. Fourteen change variables and eight domains defined by race, sex, and father's education were chosen for the analysis. Variance components and efficiency were calculated for each of the 238 statistics defined by the above variables and domains, for each of the nine sample designs. The average efficiencies of each design were calculated for all of the statistics, national statistics, and domain statistics. The optimal designs--those with the highest efficiencies--were defined as follows: (1) for national statistics, 12 students sampled per school; (2) 15 white students per school; 18 for blacks, other races, males, and females; and 21-27 for the domains defined by father's occupation; and (3) for all statistics, the current NLS design of 18 students per school. Designs which sampled nine to 27 students per school all had high efficiencies for national and domain statistics. Base-year NLS and first followup questions used to define change variables are appended. (Author/GDC)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED167571

22U-884-3

EFFICIENCY STUDY OF THE NLS SAMPLE DESIGN

by

R. P. Moore
B. V. Shah

BEST COPY AVAILABLE

Statistics Research Division
Research Triangle Institute

October 1975

TM007 681

NLS LIBRARY

RESEARCH TRIANGLE PARK, NORTH CAROLINA 27709

TABLE OF CONTENTS

		Page
LIST OF TABLES		iii
I.	INTRODUCTION	1
II.	VARIABLES	3
III.	COST MODEL	7
	A. Distance Traveled Model	7
	B. Cost of A Single Follow-Up Survey	10
	C. Costs of All Surveys in An NLS Cohort	12
IV.	RESULTS	17
	A. Efficiencies	17
	B. Variance Component Ratios	22
	C. Comparisons with Base-Year Efficiency Study	22
V.	SUMMARY	25
REFERENCES		27
APPENDIX A		28

LIST OF TABLES

Page

1 - Estimated Direct Charges Allocated to School and Student Costs, NLS Follow-Up Survey 14

2 - Estimated Man-Months of Direct Labor, by Task, by Position, NLS Follow-Up Survey 14

3 - Calculation of Estimated Variable Cost Parameters, C_1 and C_2, for an NLS Follow-Up Survey 15

4 - Specification of Nine Designs with Approximately Equal Cost. . . . 16

5 - Mean Efficiencies by Type of Statistic 19

6 - Standard Deviations of Efficiencies by Type of Statistic 20

7 - Mean Efficiencies for Domain Means and Differences between Domain and National Means 21

8 - Number of Respondents and Ratios of Variance Components, by Domains. 23

I. INTRODUCTION

Looking ahead to the NLS New Cohort Survey, this design efficiency study seeks to determine the "optimum" number of schools and seniors per school to select in a two-stage design so as to minimize the variance of New Cohort statistics subject to likely budget constraints. The optimal numbers of schools and seniors per school were determined previously. The results of this previous study were based on variance components and cost modeling for a single-cost survey. But the follow-up surveys of an NLS cohort involve different cost relationships, and possibly different variance ratios, than does a single one-time survey. Thus, determination of an "optimal" two-stage sample design for an NLS cohort should be based on cost modeling which reflects the costs of a base-year survey and several follow-up surveys. Similarly, the variance component estimates for a longitudinal survey should be calculated for variables describing changes in student plans, attitudes, and activities over time. Thus, the variables used in this analysis were defined as changes (or lack of changes) in the individuals' responses from the Base-Year (BY) survey to the First Follow-Up (1FU) survey.

Although it might have been more appropriate to use variables involving changes over a longer span of time, one-year change data were the only data available. Comparing the optimum design determined from one-year change variables with that determined using only Base-Year variables may give an indication of the nature of the optimum design for the entire project. We were searching for an efficient design for the entire system of surveys--the Base-Year survey of seniors plus five Follow-Up surveys conducted one, two four, six, and eight years thereafter.

As with the Base-Year design optimization, it was assumed possible to specify a number of equal-cost designs for the NLS and that the criterion of average efficiency of a given design over many statistics was a reasonable criterion for ranking competing designs. In order to optimize the design for the entire system of NLS surveys, the cost model was written as the total variable cost associated with a Base-Year survey and five Follow-Up surveys. The cost of conducting a single Follow-Up survey was modeled and five times that Follow-Up survey cost was added to the Base-Year cost components estimated during the Base-Year variance components project.

Fourteen response variables were defined reflecting change (or lack of change) in activity state, attitude, or experience from the Base-Year to the First Follow-Up survey. Variance components for these change variables were estimated using Seeger's formulae and Taylorized deviations [2,3]. The "efficiency" of each of nine equal-cost sample designs was estimated for each of 238 statistics using the method of Shah and Folsom [4]. These 238 statistics consist of a national mean, eight subgroup means, and eight subgroup mean versus national mean differences for each of the 14 change variables. The average efficiency of each design was calculated for all statistics, national statistics, and domain statistics. The "optimum" design for each type of statistic was then chosen as the design with the highest average efficiency.

Since the data file used for this study was constructed prior to the availability of all First Follow-Up data, it included only 15,578 matched cases rather than the 15,635 sample students who completed both Base-Year and First Follow-Up Student Questionnaires. Base-Year adjusted student weights calculated by ETS for 17,726 Base-Year participants were used in this analysis. Some students completed both questionnaires but did not respond to some items sufficiently to determine values for some of the change variables. The "VARCOMP" program assumes missing values occur in the sample at random. The statistics themselves are calculated based on weighted sums for the respondents, but the estimated variance components include contributions to the variance due to nonresponse and missing data values.

II. VARIABLES

A total of 14 "across time" or change variables, nine binomial and five multinomial, were chosen for the efficiency analysis. The variables were defined in terms of the correspondence of individual student's abilities, attitudes, and activities for the two time points. Variables one through five were defined as differences between composite scores computed from responses to several questions asked at each time point. For variables six through thirteen, a change in expectations, aspirations, goals, or plans was coded "1" and no change was coded "0." Variable 14 was a simple proportion calculated using data from both time periods. The formulation of the 14 change variables is described in the following paragraphs. The particular NLS questionnaire items referred to in this section are shown in appendix A.

1. Change in self-concept composite - Calculated for each respondent as the sum of 1FU question 15, items a, c, d, and h minus the sum of BY question 21, items a, c, d, and h (after recoding the responses to each item to: agree strongly = 1, agree = 2, no opinion = 3, disagree = 4, and disagree strongly = 5).

2. Change in control of environmental composite - Computed as the sum of 1FU question 15, items b, e, f, and g minus the sum of BY question 21, items b, e, f, and g (after recoding responses as stated above).

3. Change in work orientation composite - This variable was calculated as the sum of 1FU question 20, items a, c, and e minus the sum of BY question 20, items a, c, and e.

4. Change in community orientation composite - Computed as the sum of 1FU question 20, items f, g, and j minus the sum of BY question 20, items f, g, and j.

5. Change in family orientation composite - Computed as the sum of 1FU question 20, items b, h, and i minus the sum of BY question 20, items b, h, and i (after recoding the item-i responses to each item as: not important = 3, somewhat important = 2, and very important = 1).

6. Change in educational expectations - First, BY question 29b and 1FU question 14 responses were recoded:

BY question 29b response	1FU question 14 response	Recoded responses
1 or 2	1	1-high school
3	2 or 3	3-voc-tech
4 or 5	4 or 5	5-college program
6	6 or 7	6-post graduate

The change variable was then defined by assigning each student a value of "1" if his educational expectations changed, or the value "0" otherwise.

7. Change in educational aspirations - First, responses to BY question 29A and 1FU question 12 were recoded as described above. Next, the change variable for each student was defined as "1" if his educational aspirations had changed from the BY to 1FU, or "0" otherwise.

8. Change in type of community of respondent's residence - BY question 95 and 1FU question 6A responses were compared and students were assigned values of "1" if they reported a change in type of community, or "0" otherwise.

9. Change in occupational goals - Student responses to BY question 25A and to 1FU question 19 were recoded as follows:

BY question 25A response	1FU question 19 response	Recoded responses
1	1	1-clerical
2	2	2-craftsman
3	3	3-farm
4	4	4-homemaker
5	5	5-laborer
6	6	6-manager
7	7	7-military
8	8	8-operative
9	9,10,14	9-professional
10	11	10-proprietor
11	12	11-protective service
12	13	12-sales
13	15	13-service
14	16	14-technical

This change variable was then defined by assigning each student a value of "1" if his occupational goals had changed, or "0" otherwise.

10. Correspondence of educational expectations in BY with educational activities in 1FU - First, responses were recoded as follows:

BY question 29B response	1FU question 26B response	Recoded responses
1,2,3	1,99	1-voc-tech or none
4	2	2-junior college
5,6	3	3-four-year college

This change variable was constructed by assigning each student a value of "0" if his BY educational expectations agreed with his 1FU educational activities, or "1" otherwise.

11. Correspondence of BY financial aid plans with 1FU financial aid status - This variable was defined in terms of plans and outcomes involving use or nonuse of three categories of aid--federal scholarships, federal loans, and nonfederal scholarships or loans. First, six preliminary (0,1) variables were defined for each student as follows:

From BY question 22:

I_1 = { 1 if response of "Yes, I plan to use it" was obtained for any of the items d, e, g, j, l, n, p, or q, or
0 otherwise

I_2 = { 1 if response of "Yes, I plan to use it" was obtained for any of the items h, i, k, or m, or
0 otherwise

I_3 { 1 if response of "Yes, I plan to use it" was obtained for any of the items a, b, c, or r, or
0 otherwise.

From 1FU question 47:

J_1 = { 1 if amounts were entered for any of the codes 07, 08, 10, 11, 12, 21, 22, 23, 25, or
0 otherwise

J_2 = { 1 if amounts were entered for any of the codes 15, 18, 19, or 20, or
0 otherwise.

J_3 = { 1 if amounts were entered for any of the codes 09, 13, 14, 16, or 17, or
0 otherwise.

Finally, the value of the change variable was assigned as "0" if $I_1 = J_1$, $I_2 = J_2$, and $I_3 = J_3$. If one or more of the I,J pairs had different values, the change variable was assigned the value "1" which indicated the students' BY financial aid plans were not in agreement with his 1FU financial aid status.

12. Correspondence of planned activity state in BY with actual activity state in 1FU - Planned activity states obtained as responses to BY question 31 were recoded as follows:

BY question 31 response	Planned activity state
1,2,9	1-work
5,7	2-voc-tech
6,8	3-college
3	4-military
4	5-homemaker
10, blank	6-other

Actual activity states in 1FU were based on the responses to 1FU question 1. Since respondents could indicate more than one activity, a sequential scheme was used to assign a unique activity state code to each individual. The priority order of responses was d, b, c, a, e, and f. That is, on active duty in military (d) response established the student's activity state as military regardless of any other items he may have checked. Then the change variable was assigned a value for each student as "0" if the BY planned activity state and 1FU actual activity state agreed, or "1" if they were different.

13. Correspondence of planned activity state in BY with 1FU planned activity state for 1974 - The planned activity state described in question 12 above was also used to define this analysis variable. The planned 1974 activity state was defined based on the responses to 1FU question 16 items using essentially the same approach as for 1FU question 1 responses above. That is, respondents could indicate more than one activity, so a sequential scheme was used to assign a unique planned activity code to each respondent. The priority order of responses to 1FU question 16 was d, b, c, a, and e. Then the change variable was constructed by assigning the value "0" to each student whose BY planned activity state and 1FU planned 1974 activity state agreed. Otherwise the change variable was assigned a "1" value.

14. Proportion of students who have above average ability and who are enrolled in higher education - Students were assigned the value "1" if they scored at or above the median BY test battery score and if they were enrolled in higher education as indicated by a response of "applies to me" for 1FU question 1b or 1c. Students were assigned the value "0" if they scored below the median test score or if they did not check "applies to me" for either 1FU item 1b or 1c.

III. COST MODEL

The cost model for the design efficiency analysis had to include the total variable costs incurred in conducting a Base-Year survey and five Follow-Up surveys. Modeling of the distance traveled by interviewers in the course of an NLS Follow-Up survey is described in section A, below, as a basis for allocating the total variable costs of an NLS Follow-Up survey as school and student costs (section B). The costs of conducting an NLS Base-Year survey had been modeled in a previous report [1]. The cost parameters for a linear model of the total variable cost of an entire NLS cohort (Base-Year and five Follow-Up surveys) are estimated in paragraph C, below. Nine two-stage designs with approximately equal costs are also specified.

A. Distance Traveled Model

The clustering of sample students in groups of 18 per school reduces the interviewer's travel cost for follow-up visits with mail questionnaire nonrespondents. The cost reduction is dampened, however, by the fact that the population is quite mobile. Since interviewers generally spend more time traveling and locating the nonrespondents than in actually interviewing them, the distance model may be used to predict interviewer salary costs, as well as travel costs.

The average distance between k points, assumed to be uniformly scattered, throughout an area A, is $\bar{D} = \sqrt{\frac{A}{k}}$. A simple rational for this result can be obtained by considering a square figure with sides of length $\sqrt{A}$ covered by a square grid of k equally spaced points. A more general motivation is provided by Beardwood, et.al.[2]. The total distance traveled to visit all k points once is $D = \sqrt{kA}$. In our case we assumed $D = D_1 + D_2 + D_3$ where

D_1 = distance traveled to visit all sample school locations,

D_2 = distance to visit all nonrespondents who still live within the county of the school location, and

D_3 = distance traveled to visit all nonrespondents who have moved from the county of the school location.

Assume a working staff of I interviewers, each of whom had to visit n_1/I school locations within his work area of area A/I where

n_1 = total number of sample schools, and

A = area of United States.

Thus we write D_1 as

$$D_1 = I\sqrt{\frac{n_1}{I}\ \frac{A}{I}} = \sqrt{n_1 A}\ .$$

Next assume that in each school location area, the interviewer visits $(1 - r_m)\ (1 - m)\ n_2$ nonrespondents who still reside in the area, where

r_m = proportion of the sample students who return a completed questionnaire by mail,

m = proportion of sample students who moved out of the county their school is located in, and

n_2 = number of sample students per school.

The nonmovers are assumed to be equally spaced across an average sized county of area in square miles equal to A divided by 3,141, the number of continental U.S. counties.

Now we can write D_2 in our model as

$$D_2 = n_1\ \sqrt{(1 - r_m)(1 - m)\ n_2\ A/3{,}141}\ ,$$

the total distance traveled by all interviewers to visit all nonmover nonrespondents once the interviewers have reached the area of the school location.

Finally, assume a separate task requires each interviewer to visit a total of $(1 - r_m)m\ n_1\ n_2/I$ movers, who are assumed to be equally spaced throughout the interviewer's work area, of average size A/I, so that

$$D_3 = I\sqrt{\frac{(1 - r_m)m\ n_1\ n_2}{I}\ \frac{A}{I}} = \sqrt{(1 - r_m)m\ n_1\ n_2\ A}\ .$$

Now the total distance traveled in miles, under the assumption of our model was written as $D = D_1 + D_2 + D_3$ or

$$D = \sqrt{n_1 A} + n_1\ \sqrt{(1 - r_m)\ (1 - m) n_2\ A/3{,}141} + \sqrt{(1 - r_m)m\ n_1 n_2 A}\ .$$

First Follow-Up (1FU) response experience indicates a mail questionnaire response rate in the neighborhood of .60. An approximate value for the fraction of movers, namely m = .4, was obtained by noting that of the 23,020 students who were mailed First Follow-Up questionnaires, roughly 60 percent had updated mailing addresses in the same zip code area as their parents. This estimate of movers will be biased upward by students moving to different zip code areas in the same county and downward by those who have moved from the county but retained their parents' mailing address. In a forthcoming report, an improved estimate of mail nonrespondent mobility will be obtained using 1FU personal interview data. The impact of any significant deviation from m = .4 on the deviation of "optimum" designs will be explored in this later report.

Finally, substituting the values $r_m = .60$, m = .40, and the continental U.S. land area, A = 3,615,122 square miles, the equation for D becomes

$$D = 1{,}901\sqrt{n_1} + 16.62\, n_1\sqrt{n_2} + 761\sqrt{n_1 n_2} \quad . \qquad (1)$$

Fixing $n_1 n_2$ as 21,600 and letting n_1 vary, the equation was rewritten as

$$D = 4{,}344\sqrt{n_1} + 111{,}844$$

which agrees quite closely with the linear model

$$D = 182{,}000 + 65\, n_1$$

within the range of n_1 we were interested in, $600 \leq n_1 \leq 1{,}800$. For student sample sizes $n_1 n_2$ in the neighborhood of 21,600, say $\pm$ 1,000,

$$D = n_1 65 + n_1 n_2\, 8.426$$

is a good approximation for the general form of D in equation (1). With the assumption that interviewer salaries and travel costs are proportional to the distance traveled-D, the fraction of these costs that are school-related for given values of n_1 and n_2 is therefore

$$F(n_1 n_2) = n_1 65/[n_1 65 + n_1 n_2\, 8.426] \quad . \qquad (2)$$

substituting the current design parameters $n_1 = 1,200$ and $n_2 = 18$ into equation (2) one observes that approximately 30 percent of the associated interviewer costs should be assigned to the school component with the remainder going to the student component.

While a more complex model might better reflect the real world, the distance model was considered an adequate representation of how these costs varied with different degrees of sample student clustering. Of course, interviewers were not always able to travel the shortest routes, respondents were not equally spaced geographically, and one or more callbacks were necessary in order to locate some of the students. If we assume the interviewers had to travel a total distance of $2\ D = 2(182,000 + 65\ n_1) = 520,000$ miles to complete the tasks, these model results are within range of the 600,000 miles that were budgeted for this task.

B. Cost of a Single Follow-Up Survey

A linear cost model of the form

$$C = C_0 + n_1 C_1 + n_1 n_2 C_2$$

was used to model the cost of the NLS second Follow-Up survey where

- C = total cost of the second Follow-Up survey,
- C_0 = fixed cost which did not depend on the sample design,
- C_1 = marginal cost of adding an additional sample school,
- C_2 = marginal cost of adding an additional student to the sample in a school,
- n_1 = number of sample schools, and
- n_2 = number of sample students per school.

The cost modeling was based on estimates provided by RTI's second Follow-Up proposal.* These estimates were the best available source of information about the total cost of conducting an NLS Follow-Up survey. For the sample design optimization, it was necessary to isolate the variable cost portion of a Follow-Up survey, and to estimate C_1 and C_2, the vital cost parameters

*RTI Proposal No. 22-74-16-01 dated February 27, 1974.

of the model.

There were ten major tasks in the 1972 NLS second Follow-Up project:

Task 1 - general project direction,

Task 2 - implementation of tracing procedures,

Task 3 - procurement, distribution, and receipt control of survey questionnaires,

Task 4 - development of critical data base,

Task 5 - computer software support, further development of the automated survey system,

Task 6 - reliability and validity studies,

Task 7 - field interviews of nonrespondents,

Task 8 - analysis of data,

Task 9 - preparation of special interest reports, and

Task 10 - design, preparation, and field test for third Follow-Up NLS class of 1972.

Tasks 1, 4, 5, 6, 8, 9, and 10 were assumed to involve only fixed costs, within the range of two-stage sample design options considered appropriate for the NLS project. Thus, the variable portion of the total survey cost was assumed to be associated with some of the costs necessary to perform tasks 2, 3, and 7.

Virtually all of the direct charges estimated for tasks 2, 3, and 7 were regarded as variable costs. Exceptions were some of the reproduction costs from tasks 2 and 7, and the central staff travel costs from task 3. The majority of the costs were considered to occur in proportion to the total number of students in the sample and the costs were assumed to be independent of the number of sample schools. Central staff travel expenses for hiring, training, and supervising field interviewers occur in proportion to the number of interviewers, which in the case of the NLS follow-ups was considered to be proportional to the total number of sample students. The interviewer salaries and mileage expenses were assumed to partially depend on the number of sample schools (n_1) and these expenses were allocated 30 percent to school-related costs and 70 percent to student-related costs, based on the distance modeling described in section A above. Table 1 shows the dollar costs for direct charges, by cost item for school- and student-related cost categories.

The portion of the direct technical labor budgeted for tasks 2, 3, and 7 and considered to be variable, rather than fixed, cost is shown in table 2, by task. The time budgeted for task direction, computer programming, secretarial support, designing procedures, and preparing manuals were regarded as fixed cost items and are not shown. The direct labor considered to be variable in nature is also shown separately in table 2 for Survey Specialists, Research Assistants, and Survey Clerks. All of the variable cost for central staff direct labor was assumed to be student-related cost, rather than school-related cost.

The cost parameters, C_1 and C_2, were estimated based on the data from tables 1 and 2. The calculations are shown in table 3. Direct labor was converted to dollars using average cost factors of $1,080 per Survey Specialist man-month, $730 per Research Assistant man-month, and $550 per Survey Clerk man-month. Other direct costs estimated in table 1 were used. Overhead or indirect costs were included using RTI's currently approved proposal rates for indirect technical expense, and overall management and service expense. An approximate average fixed fee of five percent was also included to arrive at the total estimated variable costs of $56,993 for school costs and $729,816 for student costs. Dividing the school cost total by 1,200 and the student cost estimate by 21,600 gave the estimated cost parameters, $C_1 = \$47.49$ and $C_2 = \$33.79$ shown in table 3.

C. Costs of All Surveys in an NLS Cohort

In order to attempt to optimize the sample design for all surveys of an NLS cohort, the cost parameters in table 3 were assumed representative for each of five Follow-Up surveys and the Base-Year cost parameters developed earlier were used for the Base-Year survey cost. A linear cost model was then developed to represent the cost of an NLS Base-Year survey and five Follow-Up surveys. The school component of cost was determined as $C_1 = \$130.38 + 5\ (\$47.49) = \$367.83$ while the student component was computed as $C_2 = \$11.28 + 5\ (\$33.79) = \$180.23$. This uses the Base-Year components $C_1 = \$130.38$ and $C_2 = \$11.28$ determined in the earlier paper (see table 3 of reference 1) and the Follow-Up survey cost components shown in table 3.

A set of nine cost equivalent two-stage sample designs, in terms of the total variable cost, $C-C_0$, is shown in table 4. The designs were obtained by varying n_2, the number of students per school, and solving for the n_1, the number of sample schools, which would equalize the total variable cost, $C-C_0$.

Table 1 - Estimated Direct Charges Allocated to School and Student Costs, NLS Follow-Up Survey

Item	School cost (C_1) dollars	Student cost (C_2) dollars
Printing subcontract	-	86,940
Other reproductions	-	1,750
Materials	-	4,150
Communications and shipping	-	88,730
Services	-	4,225
Temporary labor	-	104,089
Staff travel	-	40,500
Interviewer salaries	27,300	63,700
Interviewer mileage	21,600	50,400
Total direct charges	48,900	444,484

Table 2 - Estimated Man-Months of Direct Labor, by Task, by Position, NLS Follow-Up Survey

Item	Survey specialists man-months	Research assistants man-months	Survey clerks man-months
Task 2	1.0	8.00	1.00
Task 3	5.0	30.00	16.25
Task 7	45.0	9.0	9.0
Total	51.0	47.0	26.25

Table 3 - Calculation of Estimated Variable Cost Parameters, C_1 and C_2, for an NLS Follow-Up Survey

Item	School cost (C_1) dollars	Student cost (C_2) dollars
Direct labor		
Survey specialists	-	55,080
Research assistants	-	34,310
Survey clerks	-	14,438
Total		103,828
Indirect technical expense[1/]	-	77,871
Direct charges	48,900	444,484
Overall management and service expense[2/]	5,379	68,880
Total estimated costs	54,279	695,063
Fixed fee[3/]	2,714	34,753
Total estimated cost plus fixed fee	56,993	729,816
Estimated cost parameters	47.49	33.79

[1/] Computed as 75 percent of direct labor

[2/] Computed as 11 percent of direct labor, indirect technical expense, and direct charges

[3/] Calculated as five percent of total estimated costs.

Table 4 - Specification of Nine Designs with Approximately Equal Cost*

Design	Number of schools (n_1)	Number of students per school (n_2)	Total No. of students ($n_1 n_2$)	Total variable cost ($C-C_0$)
1	2,990	6	17,940	4,329,520
2	2,177	9	19,593	4,327,876
3	1,712	12	20,544	4,327,936
4	1,411	15	21,165	4,328,948
5**	1,200	18	21,600	4,329,600
6	1,043	21	21,903	4,326,364
7	923	24	22,152	4,327,024
8	828	27	22,356	4,328,784
9	750	30	22,500	4,326,000

*With C_1 = \$368 and C_2 = \$180.

**Design used in 1972 NLS.

IV. RESULTS

A. Efficiencies

The 14 variables described in section II were selected for the analysis of design efficiencies. For each variable, 17 sets of variance components were estimated: one set associated with the national mean, eight sets associated with domain means, and eight sets associated with differences between domain means and the national mean. (Domains or subpopulations used were males, females, whites, blacks, other races; and three levels of father's education: less than high school graduate, high school graduate, and some education beyond high school). The 238 sets of variance components were estimated using the RTI VARCOMP program [3], which utilizes Taylor series linearizations and Seeger's [4] variance components. The details of the formulae used in VARCOMP are shown in section IV of the report [1] on the NLS Base-Year efficiency study and are not repeated here.

The criterion of maximum average efficiency [5] was used for choosing the "optimal" design from a group of feasible designs with approximately equal costs. Assume there are M designs D_k, (k = 1, 2, ..., M) and N statistics Y_ℓ (ℓ = 1, 2, ..., N), and let $V(\ell,1)$ and $V(\ell,2)$ be the estimated school and student components, respectively, for Y_ℓ. Then the efficiency of design-k for statistic-ℓ is defined as

$$E(k,\ell) = \frac{\text{minimum variance for statistic-}\ell\text{ at the given cost}}{\text{variance of statistic-}\ell\text{ with design-k}}$$

where design-k has r_k sample schools and $r_k s_k$ sample students. If we let C_1 be the marginal cost of adding another school to the sample, and let C_2 be the marginal cost of adding another sample student, the numerator and denominator of $E(k,\ell)$ may be written as

$$\text{numerator} = \frac{\left(\sqrt{C_1\, V(\ell,1)} + \sqrt{C_2\, V(\ell,2)}\right)^2}{C_1\, r_k + C_2\, r_k\, s_k}$$

and

$$\text{denominator} = \frac{V(\ell,1)}{r_k} + \frac{V(\ell,2)}{r_k\, s_k} \quad ,$$

respectively. Once the efficiencies were computed, the average efficiency, M_k, and the variance of efficiencies, V_k, for each design, were computed as

$$M_k = \frac{1}{N} \sum_{\ell=1}^{N} E(k,\ell) \quad ,$$

and

$$V_k = \frac{1}{N-1} \sum_{\ell=1}^{N} \{E(k,\ell) - M_k\}^2 \quad .$$

The mean efficiencies for each design are shown in table 5, for all statistics and by type of statistic. Design 3 is the "optimal" design for national statistics; designs 5 and 6 are "optimal" for domain means and differences of domain and national means, respectively. Design 5 is "optimal" for all statistics if we are willing to give equal weight to each of the 238 statistics used in the study. If we assigned equal importance to domain and national means, then we might say that design 4 was the "optimal" design. Designs 2 through 8 each have average efficiency of .90 or higher for all types of statistics considered (see table 5). The design used for the 1972 NLS (design 5) is near the "optimal" design.

Table 6 shows the standard deviation of the efficiencies for the nine designs, by type of statistic. High mean efficiency is seen to be approximately associated with a low standard deviation.

Table 7 shows the mean efficiencies for subpopulation (domain) statistics for each design. Each of the means shown was obtained by averaging 28 efficiency estimates--14 for domain means plus 14 for differences between domain and national means. Design 4 is "optimal" for whites; design 5 is "optimal" for blacks, other races, males, and females. For the domains defined by the three levels of father's education, designs with somewhat larger cluster sizes (number of sample students per school) are "optimal"--designs 7 and 8 for students whose father did not complete high school, designs 6 and 7 for those who were graduated from high school and terminated education at that point, and design 6 for those whose fathers completed some postsecondary education. Designs 3 through 9 have average efficiencies of .90 or higher for all of the domains included in the analysis.

In summary, it appears that design 5, the current design, (18 sample students per school), is the "optimal" design considering all statistics

Table 5 - Mean Efficiencies by Type of Statistic

Design	n_2	Type of Statistic: National mean	Domain mean	Difference of domain and national mean	All
1	6	.899	.862	.840	.854
2	9	.942	.919	.903	.913
3	12	.953	.942	.932	.938
4	15	.949	.951	.946	.948
5	18	.940	.952	.952	.951
6	21	.928	.949	.953	.949
7	24	.915	.943	.951	.945
8	27	.901	.936	.948	.940
9	30	.886	.928	.944	.933
Number of efficiencies		14	112	112	238

Table 6 - Standard Deviations of Efficiencies by Type of Statistic

Design	n_2	Type of Statistic: National mean	Domain mean	Difference of domain and national mean	All
1	6	.077	.083	.081	.083
2	9	.058	.070	.072	.071
3	12	.059	.059	.062	.060
4	15	.072	.053	.053	.054
5	18	.088	.054	.047	.053
6	21	.104	.059	.044	.056
7	24	.118	.066	.045	.062
8	27	.132	.075	.049	.069
9	30	.144	.084	.054	.077
Number of efficiencies		14	112	112	238

Table 7 - Mean Efficiencies for Domain Means and Differences Between Domain and National Means

Design	n_2	Race			Sex		Father's Education		
		Whites	Blacks	Other	Males	Females	Less than high school	High school	Beyond high school
1	6	.881	.858	.865	.850	.865	.824	.826	.838
2	9	.933	.916	.927	.908	.923	.891	.889	.899
3	12	.952	.941	.955	.932	.946	.925	.920	.927
4	15	.956	.950	.967	.941	.955	.944	.935	.940
5	18	.952	.951	.970	.943	.956	.955	.942	.945
6	21	.946	.948	.969	.940	.952	.960	.944	.946
7	24	.937	.942	.965	.935	.946	.963	.944	.944
8	27	.927	.935	.959	.929	.939	.963	.942	.941
9	30	.916	.927	.953	.922	.931	.962	.940	.936
Number of efficiencies		28	28	28	28	28	28	28	28

analyzed. When one considers only a single Base-Year study [1], using cost parameters and variance components appropriate to such, larger cluster sizes are "optimal." The optimization attempted in this paper, using variance components estimated from change variables and cost parameters estimated for an entire NLS cohort of six surveys, is considered more appropriate as the basis for design decisions for the new cohort sample. In particular, if a two-stage design is used for the new cohort, it appears that a cluster size of 18 students per school is nearly "optimal."

B. Variance Component Ratios

Ratios of the school component to the sum of the school and student components of variance were calculated for each statistic. The complement, student component divided by the sum, was also calculated. Average values of these two ratios, expressed as percentages, are shown in table 8 by type of statistic. The student component for national statistics is 98.3 percent of the total, on the average. The student component accounts for a larger proportion of the sum for the smaller domains. The school component is slightly larger for the domains defined by race than for domains defined by father's education. The student component accounts for a slightly larger part of the total of the school and student component for change variables than it did for Base-Year variables.[1].

C. Comparison with Base-Year Efficiency Study

The Base-Year efficiency analysis [1] indicated "optimal" designs would involve clusters of 21 to 30 sample students per school. The results in part A of this section point to designs with clusters of from 12 to 18 students per school, depending on the relative importance assigned to national statistics and domain statistics. The average ratio of the student to school variance, $(1-\delta_1)/\delta_1$ where

$$\delta_1 = \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}$$

based on the change variables was nearly double that observed for Base-Year

Table 8 - Number of Respondents and Ratios of Variance Components, by Domains

	Average number of domain respondents	Percent of within stratum variation	
		School	Student
All students	13,508	1.7	98.3
Whites	10,560	1.2	98.8
Females	6,922	0.9	99.1
Males	6,476	0.8	99.2
Father - some education beyond high school	4,414	0.6	99.4
Father - less than high school	4,029	0.3	99.7
Father - high school graduate only	3,649	0.5	99.5
Blacks	1,513	0.8	99.2
Other races	1,199	0.7	99.3

variables. This would indicate that larger cluster sizes would be appropriate; however, the cost ratio C_1/C_2 for the entire cohort model was 2.0 whereas when only Base-Year costs were included a cost ratio of 11.0 was estimated. The combination of an increase in the variance ratio and a larger decrease in the cost ratio results in the fairly moderate decrease in optimal cluster sizes obtained from the current analysis.

As stated earlier, the "optimal" designs obtained from the current analysis, using variance components estimated from change variables and cost parameters reflecting variable costs of all surveys from an NLS cohort, are considered more appropriate as a basis for determining design parameters for the new cohort sample.

V. SUMMARY

The steps involved in the analysis of the efficiency of the NLS sample design are listed here as an aid in summarizing the study:

(A) Step 1 - A cost model was developed which expressed the total variable costs of the six surveys of an NLS cohort as a linear function of the numbers of schools and students sampled.

(B) Step 2 - A set of nine feasible sample designs with approximately equal total variable costs were derived using the cost parameters estimated in step 1.

(C) Step 3 - Fourteen change variables and eight domains defined by race, sex, and father's education were chosen for the analysis.

(D) Step 4 - Variance components were calculated using Seeger's formulae and Taylorized linearizations for each of 238 statistics defined by the variables and domains chosen for the study.

(E) Step 5 - The efficiency of each statistic ($\ell = 1, 2, \ldots, 238$) for each of the sample designs ($k = 1, 2, \ldots, 9$) was computed as the minimum variance of the statistic at the given cost divided by the variance of the statistic with design-k.

(F) Step 6 - The average efficiencies of each design were calculated for all statistics, national statistics, and domain statistics.

(G) Step 7 - The "optimum" designs for all statistics, national statistics, and domain statistics were chosen as the designs with the highest average efficiencies.

The assumptions made in the analysis are briefly listed below to aid in interpreting the results of the study:

(A) The model used to express the total variable cost of the survey was assumed to be a linear function of the numbers of schools and students sampled.

(B) The first order Taylor series linearization provided a reasonable approximation to the variance of ratio estimates and differences between ratios.

(C) The first-stage sampling units, schools, were treated as if they were selected with varying probabilities with replacement (or without replacement from a large population) for variance components purposes.

(D) Negative variance component estimates were set to zero, assuming that the associated population values were negligible.

Assumption A implies that the results might not be valid over a wide range of values for r_k and s_k, but the model should be adequate for determining cost-equivalent designs in the neighborhood of the current design.

Assumptions B, C, and D are not considered serious. The component variance model is positively biased due to ignoring the finite population correction term at the school selection stage and due to the fact that negative school component estimates were set equal to zeros. The lack of balance in the sample and the population across final strata also inflates the component estimates somewhat. Thus, the relative efficiency measures are inflated but the efficiency ranking of the competing designs is not altered.

The following points summarize the principal results of this efficiency study:

(A) The "optimal" design for national statistics involves 12 sample students per school.

(B) The "optimal" designs for domain statistics involve 15 students per school for white students; 18 per school for blacks, other races, males, and females; and from 21 to 27 per school for the domains defined by levels of father's education.

(C) The current NLS design with 18 students per school is perhaps the "optimal" overall design for all statistics.

(D) Designs with from 9 to 27 sample students per school all have high average efficiencies for both national and domain statistics.

(E) The criteria of maximum mean efficiency and minimum standard deviation of efficiencies appear to lead to the same "optimum" design choice.

(F) The ratios of the student component estimate to the sum of the school and student component estimates averaged approximately 98 percent for national statistics and 99 percent for domain statistics.

REFERENCES

[1] Moore, R. P., B. V. Shah and R. E. Folsom. "Efficiency Study of NLS Base-Year Designs," Report on RTI Project 22U-884, Research Triangle Institute, Research Triangle Park, N.C., November 1974.

[2] Beardwood, J, J. H. Holton and J. M. Hammersley. "The Shortest Path through Many Points." Proc. Cambridge Phil Soc. 55. 299-327, 1959.

[3] Shah, B. V. and C. N. Dillard. "VARCOMP Program for Computing Variance Components," Working Paper No. 30 on RTI Project 255U-796, Research Triangle Institute, Research Triangle Park, N.C., 1973.

[4] Seeger, Paul. "A Method of Estimating Variance Components in Unbalanced Designs," Technometrics 12, 1970, pp. 207-218.

[5] Shah, B. V. and R. E. Folsom. "Optimum Cluster Sizes in Multi-Purpose Sample Surveys," Proceedings of the Social Statistics Section, American Statistical Association, 1973.

[6] Levinsohn, Jay, J. A. Riccobono, and R. P. Moore. Base-Year and First Follow-Up Data File Users Manual (Preliminary), Research Triangle Institute, April 1975.

APPENDIX A

BASE-YEAR AND FIRST FOLLOW-UP QUESTIONS USED TO DEFINE CHANGE VARIABLES

21. How do you feel about each of the following statements?

(Circle one number on each line.)

		Agree strongly	Agree	Disagree	Disagree strongly	No opinion
BSYRQ21A	I take a positive attitude toward myself	1	2	3	4	5
BSYRQ21B	Good luck is more important than hard work for success	1	2	3	4	5
BSYRQ21C	I feel I am a person of worth, on an equal plane with others	1	2	3	4	5
BSYRQ21D	I am able to do things as well as most other people	1	2	3	4	5
BSYRQ21E	Every time I try to get ahead, something or somebody stops me	1	2	3	4	5
BSYRQ21F	Planning only makes a person unhappy since plans hardly ever work out anyway	1	2	3	4	5
BSYRQ21G	People who accept their condition in life are happier than those who try to change things	1	2	3	4	5
BSYRQ21H	On the whole, I'm satisfied with myself	1	2	3	4	5

15. How do you feel about each of the following statements?

(Circle one number on each line.)

		Agree Strongly	Agree	Disagree	Disagree Strongly	No Opinion
FFUQ15A	I take a positive attitude toward myself	1	2	3	4	5
FFUQ15B	Good luck is more important than hard work for success	1	2	3	4	5
FFUQ15C	I feel I am a person of worth, on an equal plane with others	1	2	3	4	5
FFUQ15D	I am able to do things as well as most other people	1	2	3	4	5
FFUQ15E	Every time I try to get ahead, something or somebody stops me	1	2	3	4	5
FFUQ15F	Planning only makes a person unhappy since plans hardly ever work out anyway	1	2	3	4	5
FFUQ15G	People who accept their condition in life are happier than those who try to change things	1	2	3	4	5
FFUQ15H	On the whole, I'm satisfied with myself	1	2	3	4	5

20. How important is each of the following to you in your life?

(Circle one number on each line.)

		Not important	Somewhat important	Very important
BSYRQ20A	Being successful in my line of work	1	2	3
BSYRQ20B	Finding the right person to marry and having a happy family life	1	2	3
BSYRQ20C	Having lots of money	1	2	3
BSYRQ20D	Having strong friendships	1	2	3
BSYRQ20E	Being able to find steady work	1	2	3
BSYRQ20F	Being a leader in my community	1	2	3
BSYRQ20G	Being able to give my children better opportunities than I've had	1	2	3
BSYRQ20H	Living close to parents and relatives	1	2	3
BSYRQ20I	Getting away from this area of the country	1	2	3
BSYRQ20J	Working to correct social and economic inequalities	1	2	3

20. How important is each of the following to you in your life?

(Circle one number on each line.)

		Not important	Somewhat important	Very important
FFUQ20A	Being successful in my line of work	1	2	3
FFUQ20B	Finding the right person to marry and having a happy family life	1	2	3
FFUQ20C	Having lots of money	1	2	3
FFUQ20D	Having strong friendships	1	2	3
FFUQ20E	Being able to find steady work	1	2	3
FFUQ20F	Being a leader in my community	1	2	3
FFUQ20G	Being able to give my children better opportunities than I've had	1	2	3
FFUQ20H	Living close to parents and relatives	1	2	3
FFUQ20I	Getting away from this area of the country	1	2	3
FFUQ20J	Working to correct social and economic inequalities	1	2	3
FFUQ20K	Having leisure time to enjoy my own interests	1	2	3

29. **To answer this question, circle one number for the highest level of education you would like to attain, and also circle one for the highest level you plan to attain.**

(Circle one number in each column.)

	Would like to attain	Plan to attain
Less than high school graduation	1	1
Graduate from high school but not go beyond that	2	2
Graduate from high school and then go to a vocational, technical, business, or trade school	3	3
Go to a junior college	4	4
Go to a four-year college or university	5	5
Go to a graduate or professional school *after college*	6	6

FFUQ12 12. **How far in school would you <u>like</u> to get?**

(Circle one.)

- High school only 1
- Vocational, trade, or business school .....
  - Less than two years 2
  - Two years or more 3
- College program .....
  - Some college (including two-year degree) 4
  - Finish college (four- or five-year degree) 5
  - Master's degree or equivalent 6
  - Ph.D., M.D., or equivalent 7

FFUQ14 14. **As things stand now, how far in school do you think you actually <u>will</u> get?**

(Circle one.)

- High school only 1
- Vocational, trade, or business school .....
  - Less than two years 2
  - Two years or more 3
- College program .....
  - Some college (including two-year degree) 4
  - Finish college (four- or five-year degree) 5
  - Master's degree or equivalent 6
  - Ph.D., M.D., or equivalent 7

BSYRQ95 **95. Which best describes the location of the place in which you live?**

(Circle one.)

In a rural or farming community 1
In a small city or town of fewer than 50,000 people that is not a suburb of a larger place 2
In a medium-sized city (50,000-100,000 people) 3
In a suburb of a medium-sized city 4
In a large city (100,000-500,000 people) 5
In a suburb of a large city 6
In a very large city (over 500,000 people) 7
In a suburb of a very large city 8

FFUQ6A **6a. Which of the following best describes the location of the place where you lived in the first week of October 1973?**

(Circle one.)

In a rural or farming community 1
In a small city or town of fewer than 50,000 people that is not a suburb of a larger place 2
In a medium-sized city (50,000-100,000 people) 3
In a suburb of a medium-sized city 4
In a large city (100,000-500,000 people) 5
In a suburb of a large city 6
In a very large city (over 500,000 people) 7
In a suburb of a very large city 8

25. In the column under YOU, circle the one number that goes with the best description of the kind of work you would like to do. Under FATHER, circle the one number that best describes the work done by your father (or male guardian). Under MOTHER, circle the one number that best describes the work done by your mother (or female guardian). The exact job may not be listed but circle the one that comes closest. If either of your parents is out of work, disabled, retired, or deceased, mark the kind of work that he or she used to do.

	(Circle one number in each column.)		
	You	Father	Mother
CLERICAL such as bank teller, bookkeeper, secretary, typist, mail carrier, ticket agent	01	01	01
CRAFTSMAN such as baker, automobile mechanic, machinist, painter, plumber, telephone installer, carpenter	02	02	02
FARMER, FARM MANAGER	03	03	03
HOMEMAKER OR HOUSEWIFE	04	04	04
LABORER such as construction worker, car washer, sanitary worker, farm laborer	05	05	05
MANAGER, ADMINISTRATOR such as sales manager, office manager, school administrator, buyer, restaurant manager, government official	06	06	06
MILITARY such as career officer, enlisted man or woman in the armed forces	07	07	07
OPERATIVE such as meat cutter; assembler; machine operator; welder; taxicab, bus, or truck driver; gas station attendant	08	08	08
PROFESSIONAL such as accountant, artist, clergyman, dentist, physician, registered nurse, engineer, lawyer, librarian, teacher, writer, scientist, social worker, actor, actress	09	09	09
PROPRIETOR OR OWNER such as owner of a small business, contractor, restaurant owner	10	10	10
PROTECTIVE SERVICE such as detective, policeman or guard, sheriff, fireman	11	11	11
SALES such as salesman, sales clerk, advertising or insurance agent, real estate broker	12	12	12
SERVICE such as barber, beautician, practical nurse, private household worker, janitor, waiter	13	13	13
TECHNICAL such as draftsman, medical or dental technician, computer programmer	14	14	14

FFUQ19 19. **What kind of work will you be doing when you are 30 years old? (Circle the <u>one</u> that comes closest to what you expect to be doing.)**

CLERICAL such as bank teller, bookkeeper, secretary, typist, mail carrier, ticket agent 01

CRAFTSMAN such as baker, automobile mechanic, machinist, painter, plumber, telephone installer, carpenter 02

FARMER, FARM MANAGER 03

HOMEMAKER OR HOUSEWIFE ONLY 04

LABORER such as construction worker, car washer, sanitary worker, farm laborer 05

MANAGER, ADMINISTRATOR such as sales manager, office manager, school administrator, buyer, restaurant manager, government official 06

MILITARY such as career officer, enlisted man or woman in the Armed Forces 07

OPERATIVE such as meat cutter, assembler, machine operator, welder, taxicab, bus, or truck driver, gas station attendant 08

PROFESSIONAL such as accountant, artist, registered nurse, engineer, librarian, writer, social worker, actor, actress, athlete, politician, but <u>not</u> including public school teacher 09

PROFESSIONAL such as clergyman, dentist, physician, lawyer, scientist, college teacher 10

PROPRIETOR OR OWNER such as owner of a small business, contractor, restaurant owner 11

PROTECTIVE SERVICE such as detective, policeman or guard, sheriff, fireman 12

SALES such as salesman, sales clerk, advertising or insurance agent, real estate broker 13

SCHOOL TEACHER such as elementary or secondary 14

SERVICE such as barber, beautician, practical nurse, private household worker, janitor, waiter 15

TECHNICAL such as draftsman, medical or dental technician, computer programmer 16

NOT WORKING 17

29. To answer this question, circle one number for the highest level of education you would like to attain, and also circle one for the highest level you plan to attain.

(Circle one number in each column.)

	Would like to attain	Plan to attain
Less than high school graduation	1	1
Graduate from high school but not go beyond that	2	2
Graduate from high school and then go to a vocational, technical, business, or trade school	3	3
Go to a junior college	4	4
Go to a four-year college or university	5	5
Go to a graduate or professional school *after college*	6	6

FFUQ26B **26b. What kind of school is this?** **(Circle one.)**

Vocational, trade, business or other career training school 1
Junior or community college (two-year).....2
Four-year college or university 3
Other (please describe: _____) 4

22. Do you plan to use funds available from any of the following programs for further study beyond high school?

(Circle one number on each line.)

		No, I do not plan to use it	Yes, I plan to use it	I do not know enough about the program to answer the question
BSYRQ22A	College or university scholarship or loan (athletic or academic)	1	2	3
BSYRQ22B	State or local scholarship or loan program	1	2	3
BSYRQ22C	Scholarship from a private organization or company	1	2	3
BSYRQ22D	Veterans Administration survivors' benefits or direct benefits (GI Bill compensation or pension)	1	2	3
BSYRQ22E	ROTC Scholarship Programs	1	2	3
BSYRQ22F	College Work-Study Program	1	2	3
BSYRQ22G	Social Security Benefits for Students Age 18 to 22 (for children of retired, disabled, or deceased parents)	1	2	3
BSYRQ22H	National Defense Student Loan Program	1	2	3
BSYRQ22I	Federal Guaranteed Student Loan Program	1	2	3
BSYRQ22J	Educational Opportunity Grant Program	1	2	3
BSYRQ22K	Health Professions Student Loan Program	1	2	3
BSYRQ22L	Health Professions Scholarship Program	1	2	3
BSYRQ22M	Nursing Student Loan Program	1	2	3
BSYRQ22N	Nursing Scholarship Program	1	2	3
BSYRQ22O	General Scholarship Program	1	2	3
BSYRQ22P	Law Enforcement Education Program	1	2	3
BSYRQ22Q	Veterans Administration War Orphans Educational Assistance Program	1	2	3
BSYRQ22R	Regular bank loan	1	2	3

47. In paying for these costs, how much came from each of the following sources? (Please circle all that apply and write in the amounts.)

SAVINGS OR EARNINGS

Own savings or summer earnings ..... 01 ($_____)
College work-study programs ..... 02 ($_____)
Other earnings while taking courses ..... 03 ($_____)

INDIVIDUAL SUPPORT

Parents ..... 04 ($_____)
Husband or wife ..... 05 ($_____)
Other relatives or friends ..... 06 ($_____)

SCHOLARSHIPS OR GRANTS

Basic Educational Opportunity Grant Program ..... 07 ($_____)
Supplementary Educational Opportunity Grant Program ..... 08 ($_____)
College scholarship or grant from college funds ..... 09 ($_____)
ROTC scholarship or stipend ..... 10 ($_____)
Nursing Scholarship Program ..... 11 ($_____)
Health Professions Scholarship Program ..... 12 ($_____)
State scholarships ..... 13 ($_____)
Other scholarships ..... 14 ($_____)

LOANS

Federal Guaranteed Student Loan Programs ..... 15 ($_____)
State Loan Program ..... 16 ($_____)
Regular bank loan ..... 17 ($_____)
National Defense (Direct) Student Loan Program ..... 18 ($_____)
Health Professions Student Loan Program ..... 19 ($_____)
Nursing Student Loan Program ..... 20 ($_____)

OTHER

Law Enforcement Educational Program ..... 21 ($_____)
Veterans Administration War Orphans or Survivors Benefits Programs ..... 22 ($_____)
Veterans Administration direct benefits (GI Bill compensation or pension) ..... 23 ($_____)
Vocational Rehabilitation Program benefits ..... 24 ($_____)
Social Security Benefits (for students aged 18-22 who are children of retired, disabled, or deceased parents) ..... 25 ($_____)

Source Code	Amount
FFUQ47AA	FFUQ47AB
FFUQ47BA	FFUQ47BB
FFUQ47CA	FFUQ47CB
FFUQ47DA	FFUQ47DB
FFUQ47EA	FFUQ47EB
FFUQ47FA	FFUQ47FB
FFUQ47GA	FFUQ47GB

(See manual for coding rules)

BYERQ31 **31. What is the one thing that most likely will take the largest share of your time in the year after you leave high school?**

(Circle only one number and then follow the directions beside that number.)

	(Circle one.)	Directions
Working full-time	01	→ *Go to question 32 and complete Section D.*
Entering an apprenticeship or on-the-job training program	02	→ *Skip to page 15 and complete Section E.*
Going into regular military service (or service academy)	03	→ *Skip to page 17 and complete Section F.*
Being a full-time homemaker	04	→ *Skip to page 20 and complete Section G.*
Taking vocational or technical courses at a trade or business school full-time or part time	05	→ *Skip to page 22 and complete Section H.*
Taking *academic courses* at a junior or community college full-time or part-time	06	→ *Skip to page 25 and complete Section I.*
Taking *technical or vocational subjects* at a junior or community college full-time or part-time	07	
Attending a four-year college or university full-time or part-time	08	
Working part-time, but not attending school or college	09	→ *Skip to page 29 and complete Section J.*
Other (travel, take a break, no plans)	10	→ *Skip to page 31 and complete Section K.*

1. What are you doing now?

(Circle one number on each line.)

		Applies to me	Does not apply to me
FFUQ1A	Working for pay at a full-time or part-time job	1	2
FFUQ1B	Taking vocational or technical courses at any kind of school or college (for example, vocational, trade, business, or other career training school)	1	2
FFUQ1C	Taking academic courses at a two- or four-year college	1	2
FFUQ1D	On active duty in the Armed Forces (or service academy)	1	2
FFUQ1E	Homemaker	1	2
FFUQ1F	Temporary lay-off from work, looking for work, or waiting to report to work	1	2
FFUQ1G	Other (please describe: _____)	1	2

16. What do you expect to be doing in October 1974?

(Circle one number on each line.)

		Expect to be doing	Do not expect to be doing
FFUQ16A	Working for pay at a full-time or part-time job	1	2
FFUQ16B	Taking vocational or technical courses at any kind of school or college (for example, vocational, trade, business, or other career training school)	1	2
FFUQ16C	Taking academic courses at a two- or four-year college	1	2
FFUQ16D	On active duty in the Armed Forces (or service academy)	1	2
FFUQ16E	Homemaker	1	2
FFUQ16F	Other (please describe: _____)	1	2